

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

Application of DNA Arrays to Toxicology

John C. Rockett and David J. Dix

Reproductive Toxicology Division, National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA

Docket No.: PF-0609 USN

USSN: 09/806,267

Ref. No. 5 of 11

DNA array technology makes it possible to rapidly genotype individuals or quantify the expression of thousands of genes on a single filter or glass slide, and holds enormous potential in toxicologic applications. This potential led to a U.S. Environmental Protection Agency-sponsored workshop titled "Application of Microarrays to Toxicology" on 7-8 January 1999 in Research Triangle Park, North Carolina. In addition to providing state-of-the-art information on the application of DNA or gene microarrays, the workshop catalyzed the formation of several collaborations, committees, and user's groups throughout the Research Triangle Park area and beyond. Potential application of microarrays to toxicologic research and risk assessment include genome-wide expression analyses to identify gene-expression networks and toxicant-specific signatures that can be used to define mode of action, for exposure assessment, and for environmental monitoring. Arrays may also prove useful for monitoring genetic variability and its relationship to toxicant susceptibility in human populations. **Key words:** DNA arrays, gene arrays, microarrays, toxicology. *Environ Health Perspect* 107:681-685 (1999). [Online 6 July 1999]

<http://ehpnet1.niehs.nih.gov/docs/1999/107p681-685rockett/abstract.html>

Decoding the genetic blueprint is a dream that offers manifold returns in terms of understanding how organisms develop and function in an often hostile environment. With the rapid advances in molecular biology over the last 30 years, the dream has come a step closer to reality. Molecular biologists now have the ability to elucidate the composition of any genome. Indeed, almost 20 genomes have already been sequenced and more than 60 are currently under way. Foremost among these is the Human Genome Mapping Project. However, the genomes of a number of commonly used laboratory species are also under intensive investigation, including yeast, *Arabidopsis*, maize, rice, zebra fish, mouse, rat, and dog. It is widely expected that the completion of such programs will facilitate the development of many powerful new techniques and approaches to diagnosing and treating genetically and environmentally induced diseases which afflict mankind. However, the vast amount of data being generated by genome mapping will require new high-throughput technologies to investigate the function of the millions of new genes that are being reported. Among the most widely heralded of the new functional genomics technologies are DNA arrays, which represent perhaps the most anticipated new molecular biology technique since polymerase chain reaction (PCR).

Arrays enable the study of literally thousands of genes in a single experiment. The potential importance of arrays is enormous and has been highlighted by the recent publication of an entire *Nature Genetics* supplement dedicated to the technology (1). Despite this huge surge of interest, DNA arrays are still little used and largely unproven, as demonstrated by the high ratio of review and press articles to actual data papers. Even so, the potential they offer

has driven venture capitalists into a frenzy of investment and many new companies are springing up to claim a share of this rapidly developing market.

The U.S. Environmental Protection Agency (EPA) is interested in applying DNA array technology to ongoing toxicologic studies. To learn more about the current state of the technology, the Reproductive Toxicology Division (RTD) of the National Health and Environmental Effects Research Laboratory (NHEERL; Research Triangle Park, NC) hosted a workshop on "Application of Microarrays to Toxicology" on 7-8 January 1999 in Research Triangle Park, North Carolina. The workshop was organized by David Dix, Robert Kavlock, and John Rockett of the RTD/NHEERL. Twenty-two intramural and extramural scientists from government, academia, and industry shared information, data, and opinions on the current and future applications for this exciting new technology. The workshop had more than 150 attendees, including researchers, students, and administrators from the EPA, the National Institute of Environmental Health Sciences (NIEHS), and a number of other establishments from Research Triangle Park and beyond. Presentations ranged from the technology behind array production through the sharing of actual experimental data and projections on the future importance and applications of arrays. The information contained in the workshop presentations should provide aid and insight into arrays in general and their application to toxicology in particular.

Array Elements

In the context of molecular biology, the word "array" is normally used to refer to a series of DNA or protein elements firmly attached in

a regular pattern to some kind of supportive medium. DNA array is often used interchangeably with gene array or microarray. Although not formally defined, microarray is generally used to describe the higher density arrays typically printed on glass chips. The DNA elements that make up DNA arrays can be oligonucleotides, partial gene sequences, or full-length cDNAs. Companies offering pre-made arrays that contain less than full-length clones normally use regions of the genes which are specific to that gene to prevent false positives arising through cross-hybridization. Sequence verification of cDNA clone identity is necessary because of errors in identifying specific clones from cDNA libraries and databases. Premade DNA arrays printed on membranes are currently or imminently available for human, mouse, and rat. In most cases they contain DNA sequences representing several thousand different sequence clusters or genes as delineated through the National Center for Biotechnology Information UniGene Project (2). Many of these different UniGene clusters (putative genes) are represented only by expressed sequence tags (ESTs).

Array Printing

Arrays are typically printed on one of two types of support matrix. Nylon membranes are used by most off-the-shelf array providers such as Clontech Laboratories, Inc. (Palo Alto, CA), Genome Systems, Inc. (St. Louis, MO), and Research Genetics, Inc. (Huntsville, AL). Microarrays such as those produced by Affymetrix, Inc. (Santa Clara, CA), Incyte Pharmaceuticals, Inc. (Palo Alto, CA), and many do-it-yourself (DIY) arraying groups use glass wafers or slides. Although standard microscope slides may be used, they must be preprepared to facilitate sticking of the DNA to the glass. Several different

Address correspondence to J. Rockett, Reproductive Toxicology Division (MD-72), National Health and Environmental Effects Research Laboratory, U.S. EPA, Research Triangle Park, NC 27711 USA. Telephone: (919) 541-2678. Fax: (919) 541-4017. E-mail: rockett.john@epa.gov

The authors thank R. Kavlock for envisioning the application of array technology to toxicology at the U.S. Environmental Protection Agency. We also thank T. Wall and B. Deitz for administrative assistance.

This document has been reviewed in accordance with EPA policy and approved for publication. Mention of companies, trade names, or products does not signify endorsement of such by the EPA.

Received 23 March 1999; accepted 22 April 1999.

coatings have been successfully used, including silane and lysine. The coating of slides can easily be carried out in the laboratory, but many prefer the convenience of precoated slides available from suppliers.

Once the support matrix has been prepared, the DNA elements can be applied by several methods. Affymetrix, Inc., has developed a unique photolithographic technology for attaching oligonucleotides to glass wafers. More commonly, DNA is applied by either noncontact or contact printing. Noncontact printers can use thermal, solenoid, or piezoelectric technology to spray aliquots of solution onto the support matrix and may be used to produce slide or membrane-based arrays. Cartesian Technologies, Inc. (Irvine, CA) has developed nQUAD technology for use in its PixSys printers. The system couples a syringe pump with the microsolenoid valve, a combination that provides rapid quantitative dispensing of nanoliter volumes (down to 4.2 nL) over a variable volume range. A different approach to noncontact printing uses a solid pin and ring combination (Genetic MicroSystems, Inc., Woburn, MA). This system (Figure 1) allows a broader range of sample, including cell suspensions and particulates, because the printing head cannot be blocked up in the same way as a spray nozzle. Fluid transfer is controlled in this system primarily by the pin dimensions and the force of deposition, although the nature of the support matrix and the sample will also affect transfer to some degree.

In contact printing, the pin head is dipped in the sample and then touched to the support matrix to deposit a small aliquot. Split pins were one of the first contact-printing devices to be reported and are the suggested format for DIY arrays, as described by Brown (3). Split pins are small metal pins with a precise groove cut vertically in the middle of the pin tip. In this system, 1–48 split pins are positioned in the pin-head. The split pins work by simple capillary action, not unlike a fountain pen—when the pin heads are dipped in the sample, liquid is drawn into the pin groove. A small (fixed) volume is then deposited each time the split pins are gently touched to the support matrix. Sample (100–500 pL depending on a variety of parameters) can be deposited on multiple slides before refilling is required, and array densities of > 2,500 spots/cm² may be produced. The deposit volume depends on the split size, sample fluidity, and the speed of printing. Split pins are relatively simple to produce and can be made in-house if a suitable machine shop is available. Alternatively, they can be obtained directly from companies such as TeleChem International, Inc. (Sunnyvale, CA).

Irrespective of their source, printers should be run through a preprint sequence prior to producing the actual experimental

arrays: the first 100 or so spots of a new run tend to be somewhat variable. Factors affecting spot reproducibility include slide treatment homogeneity, sample differences, and instrument errors. Other factors that come into play include clean ejection of the drop and clogging (nQUAD printing) and mechanical variations and long-term alteration in print-head surface of solid and split pins. However, with careful preparation it is possible to get a coefficient of variance for spot reproducibility below 10%.

One potential printing problem is sample carryover. Repeated washing, blotting, and drying (vacuum) of print pins between samples is normally effective at reducing sample carryover to negligible amounts. Printing should also be carried out in a controlled environment. Humidified chambers are available in which to place printers. These help prevent dust contamination and produce a uniform drying rate, which is important in determining spot size, quality, and reproducibility.

In summary, although several printing technologies are available, none are particularly outstanding and the bottom line is that they are still in a relatively early stage of evolution.

Array Hybridization

The hybridization protocol is, practically speaking, relatively straightforward and those with previous experience in blotting should have little difficulty. Array hybridizations are, in essence, reverse Southern/Northern blots—instead of applying a labeled probe to the target population of DNA/RNA, the labeled population is applied to the probe(s). With membrane-based arrays, the control and treated mRNA populations are normally converted to cDNA and labeled with isotope (e.g., ³²P) in the process. These labeled populations are then hybridized independently to parallel or serial arrays and the hybridization signal is detected with a phosphorimager. A less commonly used alternative to radioactive probes is enzymatic detection. The probe may be biotinylated, haptenylated, or have alkaline phosphatase/horseradish peroxidase attached. Hybridization is detected by enzymatic reaction yielding a color reaction (4). Differences in hybridization signals can be detected by eye or, more accurately, with the help of digital imaging and commercially available software. The labeling of the test populations for slide-based microarrays uses a slightly different approach. The probe typically consists of two samples of polyA⁺ RNA (usually from a treated and a control population) that are converted to cDNA; in the process each is labeled with a different fluor. The independently labeled probes are then mixed together and hybridized to a single microarray slide and the resulting combined fluorescent signal is scanned. After

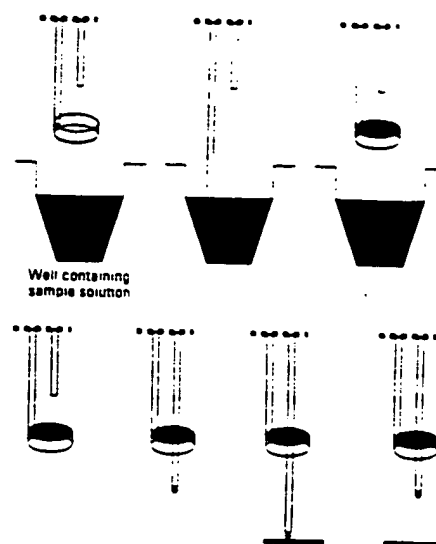


Figure 1. Genetic MicroSystems (Woburn, MA) pin ring system for printing arrays. The pin ring combination consists of a circular open ring oriented parallel to the sample solution, with a vertical pin centered over the ring. When the ring is dipped into a solution and lifted, it withdraws an aliquot of sample held by surface tension. To spot the sample, the pin is driven down through the ring and a portion of the solution is transferred to the bottom of the pin. The pin continues to move downward until the pendant drop of solution makes contact with the underlying surface. The pin is then lifted, and gravity and surface tension cause deposition of the spot onto the array. Figure from Flowers et al. (14), with permission from Genetic MicroSystems.

normalization, it is possible to determine the ratio of fluorescent signals from a single hybridization of a slide-based microarray.

cDNA derived from control and treated populations of RNA is most commonly hybridized to arrays, although subtractive hybridization or differential display reactions may also be used. Fluorophore- or radiolabeled nucleotides are directly incorporated into the cDNA in the process of converting RNA to cDNA. Alternatively, 5' end-labeled primers may be used for cDNA synthesis. These are labeled with a fluorophore for direct visualization of the hybridized array. Alternatively, biotin or a hapten may be attached to the primer, in which case fluorolabeled streptavidin or antibody must be applied before a signal can be generated. The most commonly used fluorophores at present are cyanine (Cy)3 and Cy5 (Amersham Pharmacia Biotech AB, Uppsala, Sweden). However, the relative expense of these fluorescent conjugates has driven a search for cheaper alternatives. Fluorescein, rhodamine, and Texas red have all been used, and companies such as Molecular Probes, Inc. (Eugene, OR) are developing a series of labeled nucleotides with a wide range of excitation and emission spectra which may prove to function as well as the Cy dyes.

Analysis of DNA Microarrays

Membrane-based arrays are normally analyzed on film or with a phosphorimager, whereas chip-based arrays require more specialized scanning devices. These can be divided into three main groups: the charge-coupled device camera systems, the nonconfocal laser scanners, and the confocal laser scanners. The advantages and disadvantages of each system are listed in Table 1.

Because a typical spot on a microarray can contain $> 10^8$ molecules, it is clear that a large variation in signal strength may occur. Current scanners cannot work across this many orders of magnitude (4 or 5 is more typical). However, the scanning parameters can normally be adjusted to collect more or less signal, such that two or three scans of the same array should permit the detection of rare and abundant genes.

When a microarray is scanned, the fluorescent images are captured by software normally included with the scanner. Several commercial suppliers provide additional software for quantifying array images, but the software tools are constantly evolving to meet the developing needs of researchers, and it is prudent to define one's own needs and clarify the exact capabilities of the software before its purchase. Issues that should be considered include the following:

- Can the software locate offset spots?
- Can it quantitate across irregular hybridization signals?
- Can the arrayed genes be programmed in for easy identification and location?
- Can the software connect via the Internet to databases containing further information on the gene(s) of interest?

One of the key issues raised at the workshop was the sensitivity of microarray technology. Experiments by General Scanning, Inc. (Watertown, MA), have shown that by using the Cy dyes and their scanner, signal can be detected down to levels of < 1 fluor molecule per square micrometer, which translates to detecting a rare message at approximately one copy per cell or less.

Array Applications

Although arrays are an emerging technology certain to undergo improvement and alteration, they have already been applied usefully to a number of model systems. Arrays are at their most powerful when they contain the entire genome of the species they are being used to study. For this reason, they have strong support among researchers utilizing yeast and *Caenorhabditis elegans* (5). The genomes of both of these species have been sequenced and, in the case of yeast, deposited onto arrays for examination of gene expression (6,7). With both of these species, it is relatively easy to perturb individual gene expression. Indeed, *C.*

Table 1. Advantages and disadvantages of different microarray scanning systems.

	CCD camera system	Nonconfocal laser scanner	Confocal laser scanner
Advantages	Few moving parts	Relatively simple optics	Small depth of focus reduces artifacts
	Fast scanning of bright samples	—	May have high light collection efficiency
Disadvantages	Less appropriate for dim samples	Low light collection efficiency	Small depth of focus requires scanning precision
	Optical scatter can limit performance	Background artifacts not rejected	
		Resolution typically low	

CCD, charge-coupled device.
From Kawasaki (13).

elegans knockouts can be made simply by soaking the worms in an antisense solution of the gene to be knocked out.

By a process of systematic gene disruption, it is now possible to examine the cause and effect relationships between different genes in these simple organisms. This kind of approach should help elucidate biochemical pathways and genetic control processes, deconvolute polygenic interactions, and define the architecture of the cellular network. A simple case study of how this can be achieved was presented by Butow [University of Texas Southwestern Medical Center, Dallas, TX (Figure 2)]. Although it is the phenotypic result of a single gene knockout that is being examined, the effect of such perturbation will almost always be polygenic. Polygenic interactions will become increasingly important as researchers begin to move away from single gene systems when examining the nature of toxicologic responses to external stimuli. This is especially important in toxicology because the phenotype produced by a given environmental insult is never the result of the action of a single gene; rather, it is a complex interaction of one or multiple cellular pathways. Phenomena such as quantitative trait (the continuous variation of phenotype), epistasis (the effect of alleles of one or more genes on the expression of other genes), and penetrance (proportion of individuals of a given genotype that display a particular phenotype) will become increasingly evident and important as toxicologists push toward the ultimate goal of matching the responses of individuals to different environmental stimuli.

Analysis of the transcriptome (the expression level of all the genes in a given cell population) was a use of arrays addressed by several speakers. Unfortunately, current gene nomenclature is often confusing in that single genes are allocated multiple names (usually as a result of independent discovery by different laboratories), and there was a call for standardization of gene nomenclature. Nevertheless, once a transcriptome has been assembled it can then be transferred onto arrays and used to screen any chosen system. The EPA MicroArray Consortium (EPAMAC) is assembling testes

transcriptomes for human, rat, and mouse. In a slightly different approach, Nuwaysir et al. (8) describes how the NIEHS assembled what is effectively a "toxicological transcriptome"—a library of human and mouse genes that have previously been proven or implicated in responses to toxicologic insults. Clontech Laboratories, Inc. (Palo Alto, CA), has begun a similar process by developing stress/toxicology filter arrays of rat, mouse, and human genes. Thus, rather than being tissue or cell specific, these stress/toxicology arrays can be used across a variety of model systems to look for alterations in the expression of toxicologically important genes and define the new field of toxicogenomics. The potential to identify toxicant families based on tissue- or cell-specific gene expression could revolutionize drug testing. These molecular signatures or fingerprints could not only point to the possible toxicity/carcinogenicity of newly discovered compounds (Figure 3), but also aid in elucidating their mechanism of action through identification of gene expression networks. By extension, such signatures could provide easily identifiable biomarkers to assess the degree, time, and nature of exposure.

DNA arrays are primarily a tool for examining differential gene expression in a given model. In this context they are referred to as closed systems because they lack the ability of other differential expression technologies, e.g., differential display and subtractive hybridization, to detect previously unknown genes not present on the array. This would appear to limit the power of DNA arrays to the imaginations and preconceptions of the researcher in selecting genes previously characterized and thought to be involved in the model system. However, the various genome sequencing projects have created a new category of sequence—the EST—that has partially mollified this deficiency. ESTs are cDNAs expressed in a given tissue that, although they may share some degree of sequence similarity to previously characterized genes, have not been assigned specific genetic identity. By incorporating EST clones into an array, it is possible to monitor the expression of these unknown genes. This can enable the identification of previously uncharacterized genes that may have biologic

significance in the model system. Filter arrays from Research Genetics and slide arrays from Incyte Pharmaceuticals both incorporate large numbers of ESTs from a variety of species.

A further use of microarrays is the identification of single nucleotide polymorphisms (SNPs). These genomic variations are abundant—they occur approximately every 1 kb or so—and are the basis of restriction fragment length polymorphism analysis used in forensic analysis. Affymetrix, Inc., designed chips that contain multiple repeats of the same gene sequence. Each position is present with all four possible bases. After the hybridization of the sample, the degree of hybridization to the different sequences can be measured and the exact sequence of the target gene deduced. SNPs are thought to be of vital importance in drug metabolism and toxicology. For example, single base differences in the regulatory region or active site of some genes can account for huge differences in the activity of that gene. Such SNPs are thought to explain why some people are able to metabolize certain xenobiotics better than others. Thus, arrays provide a further tool for the toxicologist investigating the nature of susceptible subpopulations and toxicologic response.

There are still many wrinkles to be ironed out before arrays become a standard tool for toxicologists. The main issues raised at the workshop by those with hands-on experience were the following:

- Expense: the cost of purchasing/contracting this technology is still too great for many individual laboratories.

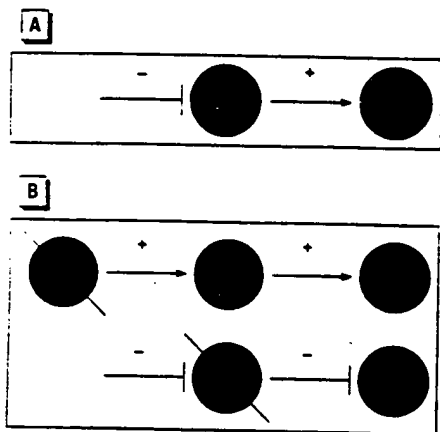


Figure 2. Potential effects of gene knockout within positively and negatively regulated gene expression networks. i_1 is limiting in wild type for expression of i_2 . (A) A simple, two-component, linear regulatory network operating on gene i_2 , where i_1 is a positive effector of i_2 and i_2 is either a positive or negative effector of i_3 . This network could be deduced by examining the consequence of (B) deleting i_1 on the expression of i_1 and i_2 , where the expression of i_2 would be decreased or increased depending on whether i_2 was a positive or negative regulator. These and other connected components of even greater complexity could be revealed by genome-wide expression analysis. From Butow (15).

- Clones: the logistics of identifying, obtaining, and maintaining a set of nonredundant, non-contaminated, sequence-verified, species/cell/tissue/tissue-specific clones.
- Use of inbred strains: where whole-organism models are being used, the use of inbred strains is important to reduce the potentially confusing effects of the individual variation typically seen in outbred populations.
- Probe: the need for relatively large amounts of RNA, which limits the type of sample (e.g., biopsy) that can be used. Also, different RNA extraction methods can give different results.
- Specificity: the ability to discriminate accurately between closely related genes (e.g., the cytochrome p450 family) and splice variants.
- Quantitation: the quantitation of gene expression using gene arrays is still open to debate. One reason for this is the different incorporation of the labeling dyes. However, the main difficulty lies in knowing what to normalize against. One option is to include a large number of so-called housekeeping genes in the array. However, the expression of these genes often change depending on the tissue and the toxicant, so it is necessary to characterize the expression of these genes in the model system before utilizing them. This is clearly not a viable option when screening multiple new compounds. A second option is to include on the array genes from a nonrelated species (e.g., a plant gene on an animal array) and to spike the probe with synthetic RNA(s) complementary to the gene(s).
- Reproducibility: this is sometimes questionable, and a figure of approximately two or three repeats was used as the minimum number required to confirm initial findings.

Again, however, most people advocated the use of Northern blots or reverse transcriptase PCR to confirm findings.

- Sensitivity: concerns were voiced about the number of target molecules that must be present in a sample for them to be detected on the array.
- Efficiency: reproducible identification of 1.5- to 2-fold differences in expression was reported, although the number of genes that undergo this level of change and remain undetected is open to debate. It is important that this level of detection be ultimately achieved because it is commonly perceived that some important transcription factors and their regulators respond at such low levels. In most cases, 3- to 5-fold was the minimum change that most were happy to accept.
- Bioinformatics: perhaps the greatest concern was how to accurately interpret the data with the greatest accuracy and efficiency. The biggest headache is trying to identify networks of gene expression that are common to different treatments or doses. The amount of data from a single experiment is huge. It may be that, in the future, several groups individually equipped with specialized software algorithms for studying their favorite genes or gene systems will be able to share the same hybridized chips. Thus, arrays could usher in a new perspective on collaboration and the sharing of data.

EPAMAC

Perhaps the main reason most scientists are unable to use array technology is the high cost involved, whether buying off-the-shelf membranes, using contract printing services, or

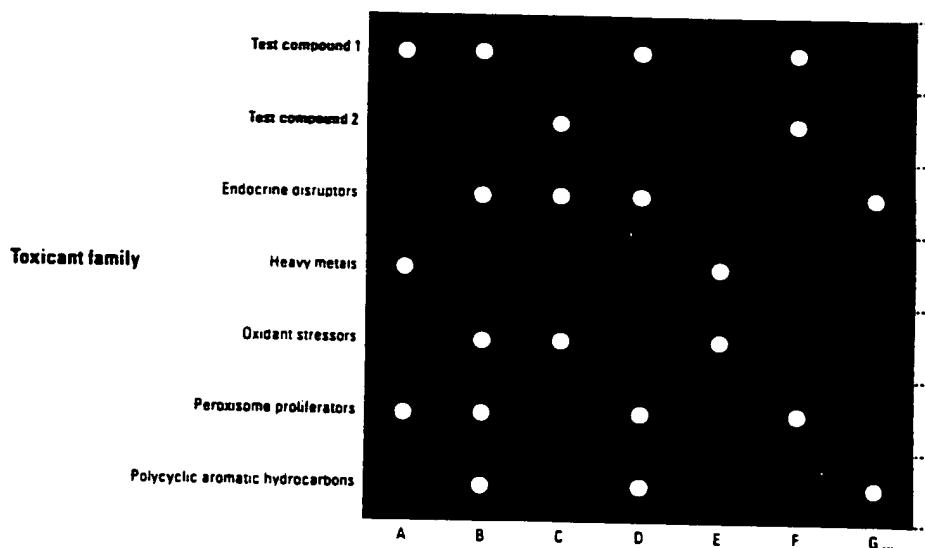


Figure 3. Gene expression profiles—also called fingerprints or signatures—of known toxicants or toxicant families may, in the future, be used to identify the potential toxicity of new drugs, etc. In this example, the genetic signature of test compound 1 is identical to that of known peroxisome proliferators, whereas that of test compound 2 does not match any known toxicant family. Based on these results, test compound 2 would be retained for further testing and test compound 1 would be eliminated.

producing chips in-house. In view of this, researchers at the RTD/NHEERL initiated the EPAMAC. This consortium brings together scientists from the EPA and a number of extramural labs with the aim of developing microarray capability through the sharing of resources and data. EPAMAC researchers are primarily interested in the developmental and toxicologic changes seen in testicular and breast tissue, and a portion of the workshop was set aside for EPAMAC members to share their ideas on how the experimental application of microarrays could facilitate their research. One of the central areas of interest to EPAMAC members is the effect of xenobiotics on male fertility and reproductive health. Of greatest concern is the effect of exposure during critical periods of development and germ cell differentiation (9), and how this may compromise sperm counts and quality following sexual maturation (10). As well as spermatogenic tissue, there is also interest in how residual mRNA found in mature sperm (11) could be used as an indicator of previous xenobiotic effects (it is easier to obtain a semen sample than a testicular biopsy). Arrays will be used to examine and compare the effect of exposure to heat and chemicals in testicular and epididymal gene expression profiles, with the aim of establishing relationships/associations between changes in developmental landmarks and the effects on sperm count and quality. Cluster, pattern, and other analysis of such data should help identify hidden relationships between genes that may reveal potential mechanisms of action and uncover roles for genes with unknown functions.

Summary

The full impact of DNA arrays may not be seen for several years, but the interest shown at this regional workshop indicates the high level of interest that they foster. Apart from educating and advertising the various technologies in this field, this workshop brought together a number of researchers from the Research Triangle Park area who are already using DNA arrays. The interest in sharing ideas and experiences led to the initiation of a Triangle array user's group.

Array technology is still in its infancy. This means that the hardware is still improving and there is no current consensus for standard procedures, quantitation, and interpretation. Consistency in spotting and scanning arrays is not yet optimized, and this is one of the most critical requirements of any experiment. In addition, one of the dark regions of array technology—strife in the courts over who owns what portions of it—has further muddled the future and is a potential barrier toward the development of consensus procedures.

Perhaps the greatest hurdle for the application of arrays is the actual interpretation of data. No specialists in bioinformatics attended the workshop, largely because they are rare and because as yet no one seems clear on the best method of approaching data analysis and interpretation. Cross-referencing results from multiple experiments (time, dose, repeats, different animals, different species) to identify commonly expressed genes is a great challenge. In most cases, we are still a long way from understanding how the expression of gene *X* is related to the expression of gene *Y*, and ordering gene expression to delineate causal relationships.

To the ordinary scientist in the typical laboratory, however, the most immediate problem is a lack of affordable instrumentation. One can purchase premade membranes at relatively affordable prices. Although these may be useful in identifying individual genes to pursue in more detail using other methods, the numbers that would be required for even a small routine toxicology experiment prohibit this as a truly viable approach. For the toxicologist, there is a need to carry out multiple experiments—dose responses, time curves, multiple animals, and repeats. Glass-based DNA arrays are most attractive in this context because they can be prepared in large batches from the same DNA source and accommodate control and treated samples on the same chip. Another problem with current off-the-shelf arrays is that they often do not contain one or more of the particular genes a group is interested in. One alternative is to obtain and/or produce a set of custom clones and have contract printing of membranes or slides carried out by a company such as Genomic Solutions, Inc. (Ann Arbor, MI). This approach

is less expensive than laying out capital for one's own entire system, although at some point it might make economic sense to print one's own arrays.

Finally, DNA arrays are currently a team effort. They are a technology that uses a wide range of skills including engineering, statistics, molecular biology, chemistry, and bioinformatics. Because most individuals are skilled in only one or perhaps two of these areas, it appears that success with arrays may be best expected by teams of collaborators consisting of individuals having each of these skills.

Those considering array applications may be amused or goaded on by the following quote from *Fortune* magazine (12):

Microprocessors have reshaped our economy, spawned vast fortunes and changed the way we live. Gene chips could be even bigger.

Although this comment may have been designed to excite the imagination rather than accurately reflect the truth, it is fair to say that the age of functional genomics is upon us. DNA arrays look set to be an important tool in this new age of biotechnology and will likely contribute answers to some of toxicology's most fundamental questions.

REFERENCES AND NOTES

1. The chipping forecast. *Nat Genet* 21(Suppl 1):3-60 (1999).
2. National Center for Biotechnology Information. The Unigene System. Available: www.ncbi.nlm.nih.gov/Schuler/UniGene [cited 22 March 1999].
3. Brown PO. The Brown Lab. Available: <http://cmgm.stanford.edu/brown> [cited 22 March 1999].
4. Chen JJ, Wu R, Yang PC, Huang JY, Sher YP, Han MH, Kao WC, Lee PJ, Chiu TF, Chang F, et al. Profiling expression patterns and isolating differentially expressed genes by cDNA microarray system with colorimetry detection. *Genomics* 51:313-324 (1998).
5. Ward S. DNA Microarray Technology to Identify Genes Controlling Spermatogenesis. Available: www.mcb.aznasa.edu/wardlab/microarray.html [cited 22 March 1999].
6. Marton MJ, DeRisi JL, Bennett HA, Iyer VR, Meyer MR, Roberts CJ, Stoughton R, Burchard J, Slade D, Dai H, et al. Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat Med* 4:1293-1301 (1998).
7. Brown PO. The Full Yeast Genome on a Chip. Available: <http://cmgm.stanford.edu/brown/yeastchip.html> [cited 22 March 1999].
8. Nuwaysir EF, Bitner M, Trent J, Barrett JC, Afshari CA. Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24(3):153-159 (1999).
9. Hecht NB. Molecular mechanisms of male germ cell differentiation. *Bioessays* 20:555-561 (1998).
10. Zacharewski TR, Timothy R. Zacharewski. Available: www.bch.msu.edu/faculty/zachar.htm [cited 22 March 1999].
11. Kramer JA, Krawetz SA. RNA in spermatozoa: implications for the alternative haploid genome. *Mol Hum Reprod* 3:473-478 (1997).
12. Stipp D. Gene chip breakthrough. *Fortune*, March 31:56-73 (1997).
13. Kawasaki E (General Scanning Instruments, Inc., Watertown, MA). Unpublished data.
14. Flowers P, Overbeck J, Mace ML Jr, Pagliughi FM, Eggers WJE, Yonkers H, Honkanen P, Montagu J, Rose SD. Development and Performance of a Novel Microarraying System Based on Surface Tension Forces. Available: <http://www.geneticmicro.com/resources/html/coldspring.html> [cited 22 March 1999].
15. Butow R (University of Texas Medical Center, Dallas, TX). Unpublished data.

SPEAKERS

Cindy Afshari NIHMS	Abdel Elkhoulou Research Genetics, Inc.
Linda Birnbaum U.S. EPA	Sue Fenton U.S. EPA
Ron Butow University of Texas Southwestern Medical Center	Norman Hecht University of Pennsylvania
Alex Chenchik Clontech Laboratories, Inc.	Pat Hurban Paradigm Genetics, Inc.
David Dix U.S. EPA	Bob Kavlock U.S. EPA
	Ernie Kawasaki General Scanning, Inc.

Steve Krawetz Wayne State University	Jim Samet U.S. EPA
Nick Mace Genetic Microsystems, Inc.	Sam Ward University of Arizona
Scott Mordecai Affymetrix, Inc.	Jeff Welch U.S. EPA
Kevin Morgan Glaxo Wellcome, Inc.	Reen Wu University of California at Davis
Elaine Poplin Research Genetics, Inc.	Tim Zacharewski Michigan State University
Don Rose Cartesian Technologies, Inc.	